

Examination of Rapidly Mutating Y-STR Loci for Increased Resolution of Common Haplotypes

Michael D. Coble, PhD*;
Becky Hill, MS; John M. Butler, PhD
Research Biologist, Applied Genetics Group

American Academy of Forensic Sciences
65th Anniversary Meeting
Washington, DC
February 22, 2013

Product Disclaimer for AAFS

- **I will mention commercial STR kit names and information, but I am in no way attempting to endorse any specific products.**
- SRM 2391c is a Standard Reference Material sold by NIST for measurement calibration purposes.
- **NIST Disclaimer**: Certain commercial equipment, instruments and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or it imply that any of the materials, instruments or equipment identified are necessarily the best available for the purpose.
- **Points of view are mine** and do not necessarily represent the official position of the National Institute of Standards and Technology or the U.S. Department of Justice. **Our group receives or has received funding from the FBI Laboratory and the National Institute of Justice.**

Outline

- Rapidly Mutating (RM) Y-STRs – Overview
- Population Genetic Parameters (current Y-STR kits)
- Utility for common Y-STR haplotypes
- Utility for close relatives
- Conclusions

What has happened in the past decade...

- **Selection of core Y-STR loci** (SWGDM Jan 2003)
- “Full” Y-chromosome sequence became available in June 2003; over 400 Y-STR loci identified (only ~20 in 2000)
- **Commercial Y-STR kits released**
 - ~~Y-PLEX 6,5,12 (2001-03)~~, **PowerPlex Y** (9/03), **Yfiler** (12/04), **PPY23** (6/12)

STR Marker Layouts for Y-STR Kits

100 bp

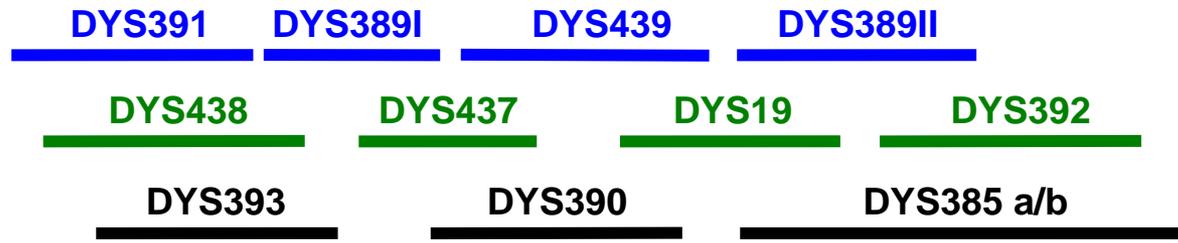
200 bp

300 bp

400 bp

2003

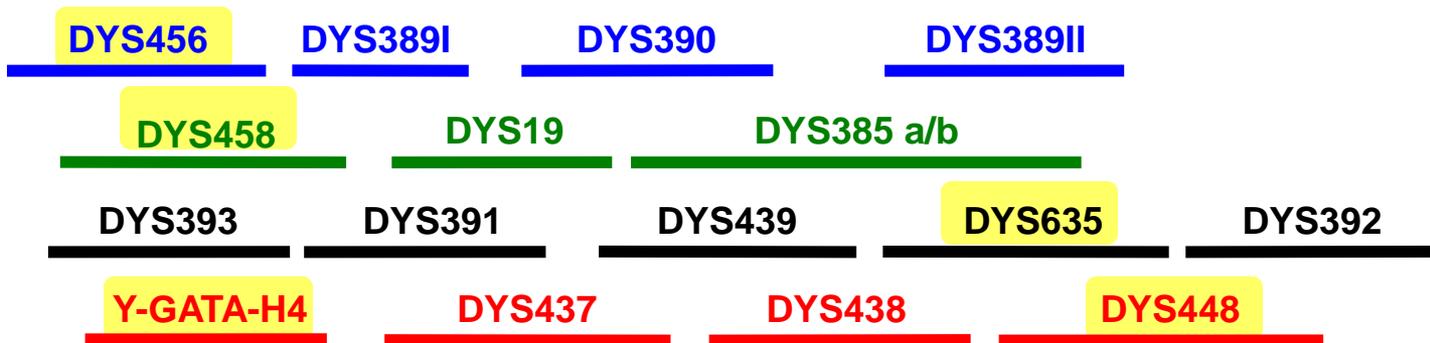
PowerPlex Y



12plex
(4-dye)

2004

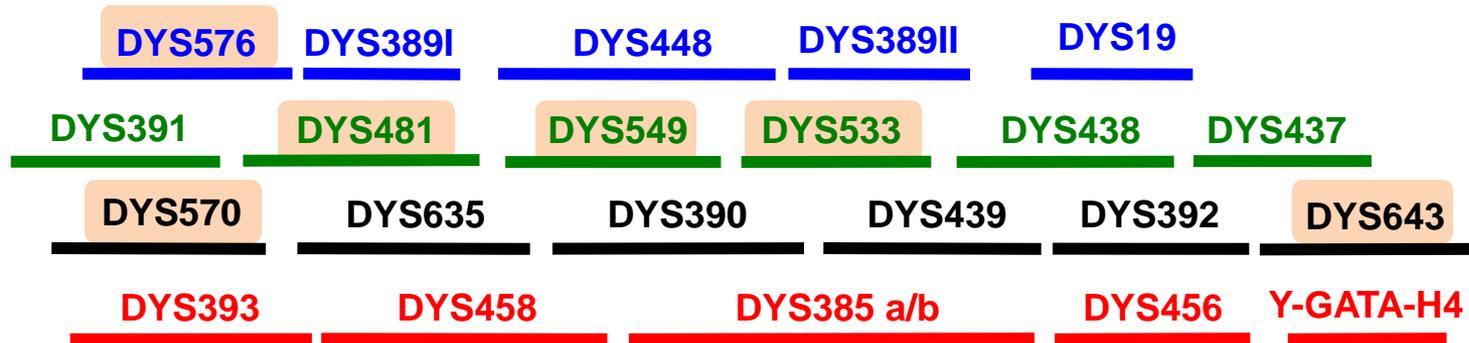
Yfiler



17plex
(5-dye)

2012

PowerPlex Y23



23plex
(5-dye)

NIST U.S. Samples (>1450)

- **NIST U.S. population samples**
 - 260 African American, 260 Caucasian, 140 Hispanic, 3 Asian
- **U.S. father/son paired samples**
 - ~**100 fathers/100 sons for each group**: 200 African American, 200 Caucasian, 200 Hispanic, 200 Asian
- **NIST SRM 2391b**, PCR-based DNA Profiling Standard (highly characterized)
 - 10 genomic DNA samples, 2 cell line samples
 - Includes 9947A and 9948
- **NIST SRM 2391c**, PCR-based DNA Profiling Standard
 - 4 genomic DNA (one mixture)
 - 2 cell lines (903 and FTA paper)

N = 1032 males

PowerPlex Y Yfiler PowerPlex Y23

haplotypes **891** **1013** **1029**

discrimination capacity 0.863 0.982 0.997

times haplotype observed PPY (12 loci) Yfiler (17 loci) PPY23 (23 loci)

1	821	998	1026
2	41	12	3
3	16	2	.
4	6	1	.
5	2	.	.
6	2	.	.
7	1	.	.
8	.	.	.
9	1	.	.
10	.	.	.
11	.	.	.
12	.	.	.
13	.	.	.
14	.	.	.
15	.	.	.
16	.	.	.
17	.	.	.
18	.	.	.
19	1	.	.

Number of unique and shared haplotypes observed with various combinations of Y-STR loci across 1032 U.S. population samples

1026 PPY23 haplotypes occur once;
and
3 sets of sample pairs cannot be resolved from one another

What has happened in the past decade...

- **Selection of core Y-STR loci** (SWGDM Jan 2003)
- “Full” Y-chromosome sequence became available in June 2003; over 400 Y-STR loci identified (only ~20 in 2000)
- **Commercial Y-STR kits released**
 - ~~Y-PLEX 6,5,12 (2001-03)~~, **PowerPlex Y** (9/03), **Yfiler** (12/04), **PPY23** (6/12)
- Many population studies performed and online databases generated with thousands of Y-STR haplotypes
- Forensic casework demonstrations showing value of Y-STR testing along with court acceptance
- Some renewed interest in Y-STRs to aid familial searching

Disadvantages of the Y-Chromosome

- Loci are not independent of one another and therefore rare random match probabilities cannot be generated with the product rule; must use haplotypes (combination of alleles observed at all tested loci)
- **Paternal lineages possess the same Y-STR haplotype** (barring mutation) and thus fathers, sons, brothers, uncles, and paternal cousins cannot be distinguished from one another
- **Not as informative as autosomal STR results**
 - **More like addition ($10 + 10 + 10 = 30$) than multiplication ($10 \times 10 \times 10 = 1,000$)**

Rapidly Mutating (RM) Y-STRs

Trying to separate
close male relatives

Rapidly Mutating Y-STRs

The American Journal of Human Genetics 87, 341–353, September 10, 2010

ARTICLE

Mutability of Y-Chromosomal Microsatellites: Rates, Characteristics, Molecular Bases, and Forensic Implications

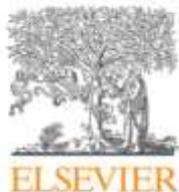
Kaye N. Ballantyne,¹ Miriam Goedbloed,¹ Rixun Fang,² Onno Schaap,¹ Oscar Lao,¹ Andreas Wollstein,^{1,3} Ying Choi,¹ Kate van Duijn,¹ Mark Vermeulen,¹ Silke Brauer,^{1,4} Ronny Decorte,⁵ Micaela Poetsch,⁶ Nicole von Wurmb-Schwarck,⁷ Peter de Knijff,⁸ Damian Labuda,⁹ H el ene V ezina,¹⁰ Hans Knoblauch,¹¹ R udiger Lessig,¹² Lutz Roewer,¹³ Rafal Ploski,¹⁴ Tadeusz Dobosz,¹⁵ Lotte Henke,¹⁶ J urgen Henke,¹⁶ Manohar R. Furtado,² and Manfred Kayser^{1,*}



Manfred Kayser

Forensic Science International: Genetics 6 (2012) 208–218

Contents lists available at ScienceDirect



Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig



13 markers
evaluated

A new future of forensic Y-chromosome analysis: Rapidly mutating Y-STRs for differentiating male relatives and paternal lineages

Kaye N. Ballantyne^{a,1,2}, Victoria Keerl^{a,1,3}, Andreas Wollstein^{a,b}, Ying Choi^a, Sofia B. Zuniga^c, Arwin Ralf^a, Mark Vermeulen^a, Peter de Knijff^c, Manfred Kayser^{a,*}

^a Department of Forensic Molecular Biology, Erasmus MC University Medical Center Rotterdam, 3000 CA Rotterdam, The Netherlands

^b Cologne Center for Genomics, University of Cologne, D-50674 Cologne, Germany

^c Department of Human Genetics, Leiden University Medical Center, 2300 RC Leiden, The Netherlands

Using Y-STRs with a higher mutation rate, father-son and brother pairs can sometimes be distinguished

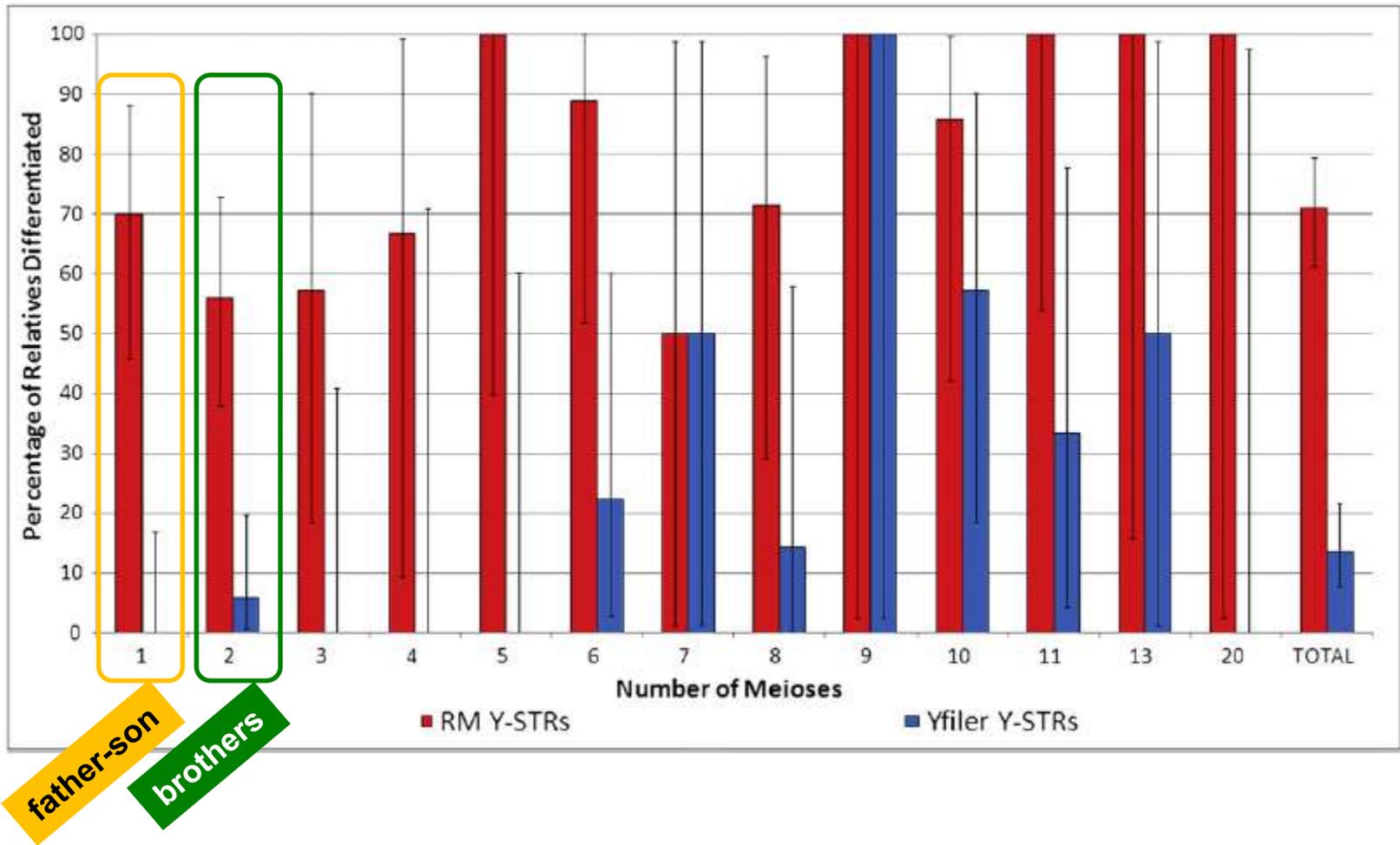
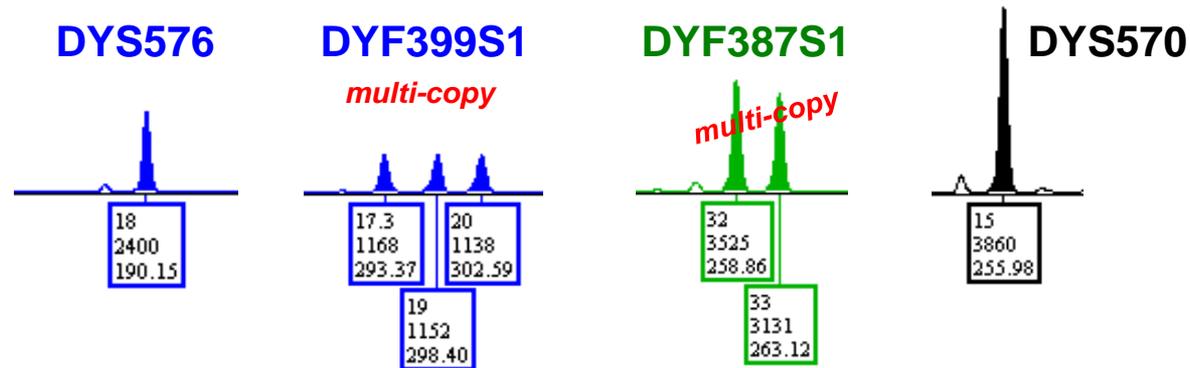


Figure 4. Male Relative Differentiation with Newly Identified 13 RM Y-STRs and Commonly Used 17 Yfiler Y-STRs
 Results from differentiating between male relatives from analyzing 103 pairs from 80 male pedigrees, sorted according to the number of generations separating pedigree members, based on 13 RM Y-STRs (in red) and 17 Yfiler Y-STRs (in blue). Error bars represent 95% binomial confidence intervals. Note that these samples are independent from the father-son pairs initially used to establish the Y-STR mutation rates.

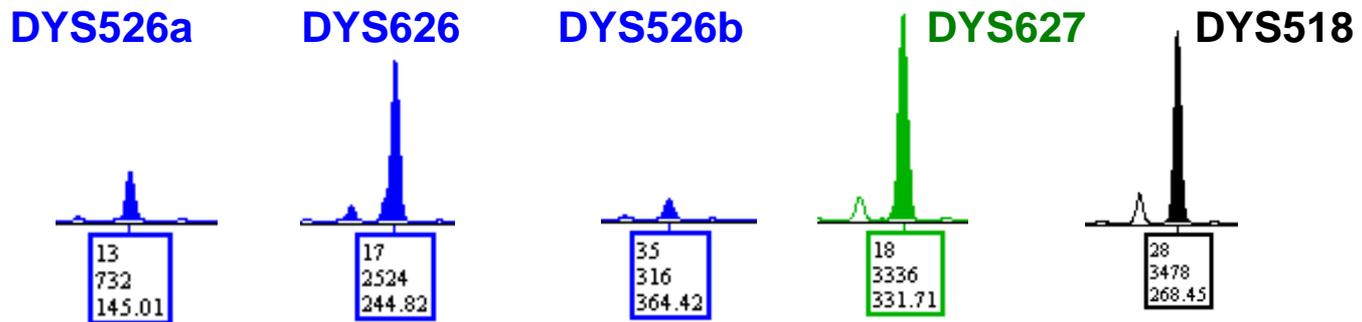
Rapidly Mutating (RM) Y-STRs

NIST supplied data from 1,296 U.S. samples (634 population + 331 father/son pairs)
to RM Y-STR Study Group led by Manfred Kayser

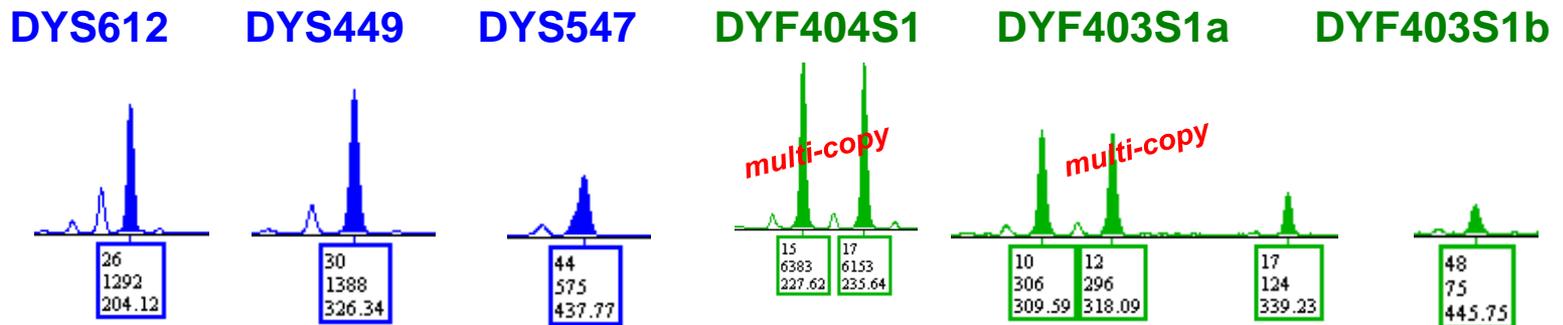
RM Y-STR
Multiplex 1



RM Y-STR
Multiplex 2



RM Y-STR
Multiplex 3



Why do these markers mutate “rapidly”?

Markers in PPY23

Locus	(average mutation rate)	“Large” number of repeats
DYS449	(1.2%)	
DYS518	(1.8%)	
DYS547	(2.4%)	
DYS570	(1.2%)	
DYS576	(1.4%)	
DYS612	(1.4%)	
DYS626	(1.2%)	
DYS627	(1.2%)	

Locus	(average mutation rate)	Multi-copy Markers
DYF387S1	(1.6%)	
DYF399S1	(7.7%)	
DYF403S1 a/b	(3.1/1.2%)	
DYF404S1	(1.3%)	
DYS526 a/b	(1.3%)	

DYS458 (0.64%) is highest in Yfiler loci where average is ~0.2%

Gene Diversity

- is a measure of the uniqueness of a particular marker in a given population

$$GD = \left(1 - \sum_i x_i^2\right)$$

↑
Relative frequency
of each allele

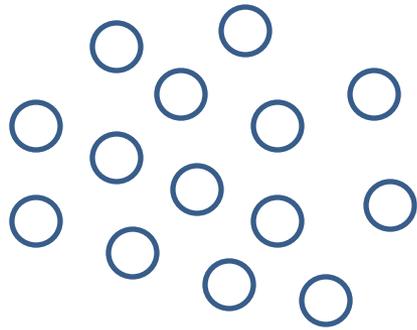
Discrimination Capacity

- is a measure of the number of unique haplotypes in a given population

$$DC = \frac{\#H}{N}$$

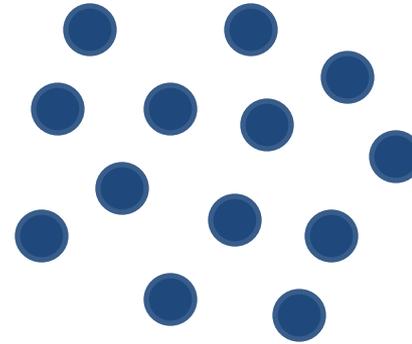
of Haplotypes

Population size



$N = 100$

Marker Y

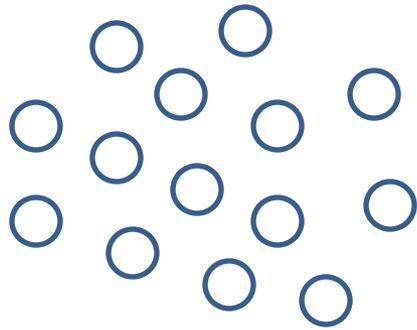


1 type = 100%

$$GD = (1 - \underbrace{\sum_i x_i^2}_{0})$$

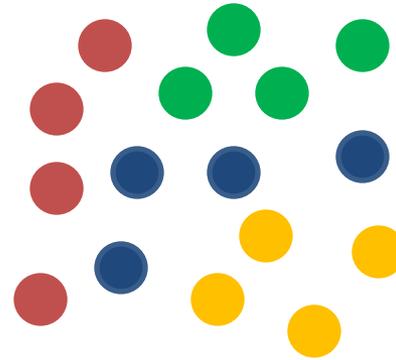
0

$$DC = 1/100 = 0.01$$



$N = 100$

Marker Y

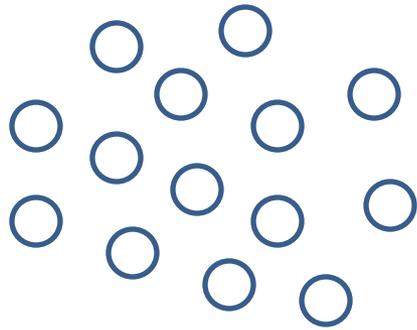


4 types = 25%

$$GD = \underbrace{\left(1 - \sum_i x_i^2\right)}$$

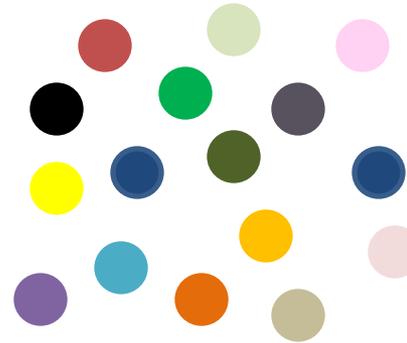
0.75

$$DC = 4/100 = 0.04$$



$N = 100$

Marker Y
→



100 types = 0%

$$GD = \underbrace{\left(1 - \sum_i x_i^2\right)}$$

0.99

$$DC = 100/100 = 1.0$$

Gene Diversity of the Markers

Marker	GD	DC
DYS576	0.766	0.035
DYF399S1	0.993	0.587
DYF387S1	0.870	0.098
DYS570	0.743	0.035
RM-01 (all)	0.9998	0.9764

Marker	GD	DC
DYS526a/b	0.923	0.138
DYS626	0.794	0.043
DYS627	0.848	0.043
DYS518	0.791	0.039
RM-02 (all)	0.9985	0.8661

DYS385a/b
GD = 0.929

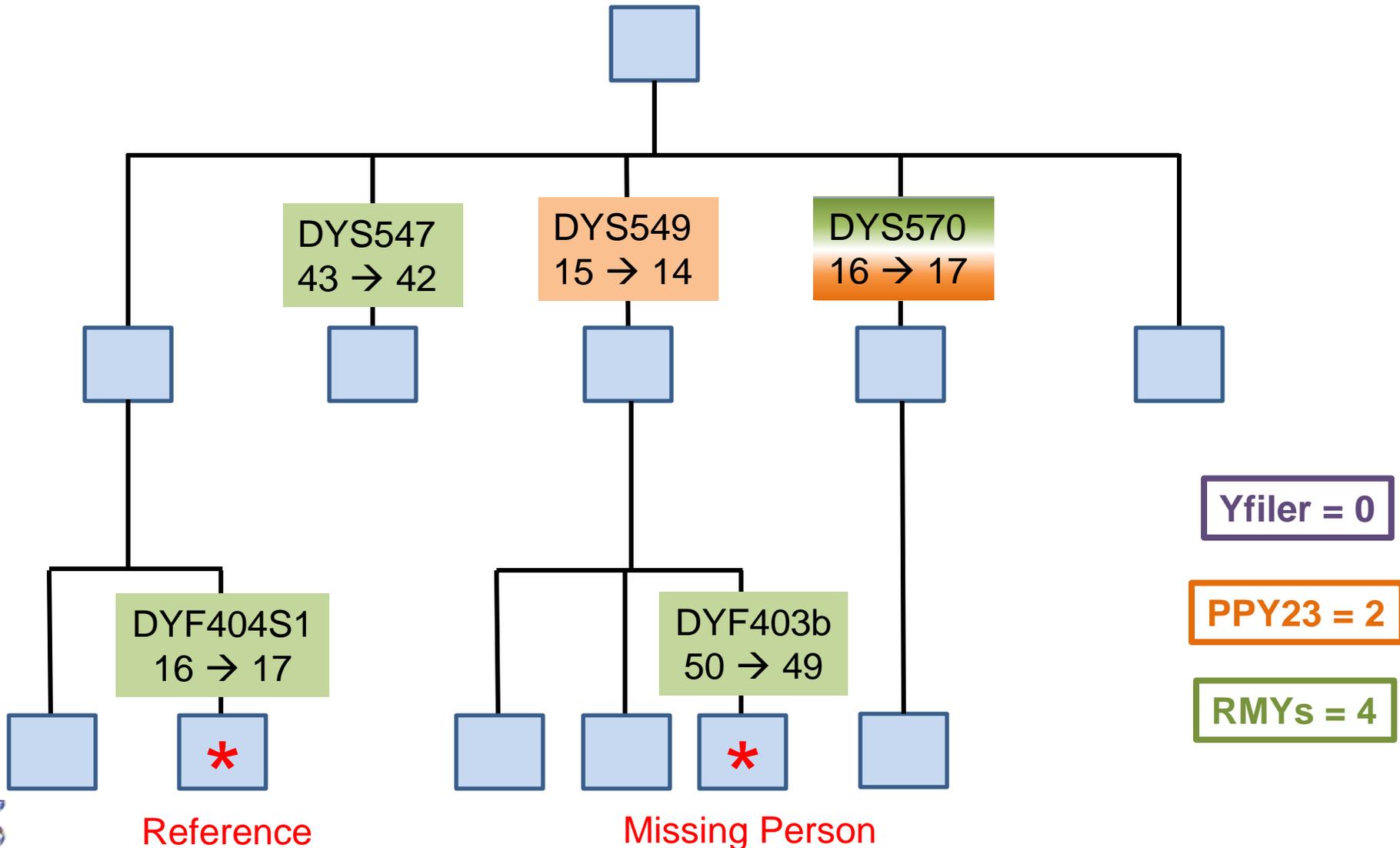
Marker	GD	DC
DYF403S1a/b	0.923	0.791
DYF404S1	0.902	0.110
DYS612	0.832	0.043
DYS449	0.796	0.043
DYS547	0.798	0.039
RM-03 (all)	1.000	0.9984

Resolution of
630/631 haplotypes

Paternal Relatives in the Database

	PPY-23	mtDNA	Kinship Index	RM Mutations
Y27	match	n/a	Father-Son	0
Y28	match	n/a	254,325,532	
Y16	match	match	Full Sib	0
Y17	match	match	155,463	
ZT79994	match	match	Full Sib	1
ZT79995	match	match	56,327	

Y-STR mutations in a Paternal Lineage



Mutation Rate Information

Meioses	Mutations	Group
63	15	AfAm
89	25	Asian
91	11	Caucasian
88	20	Hispanic
331	71	total
		(21.4%)

+1 Repeat (Son)	-1 Repeat (Son)	
8	6	AfAm
11	13	Asian
5	6	Caucasian
8	12	Hispanic
+2 Repeat (Son)	-2 Repeat (Son)	
0	1	AfAm
1	0	Asian
0	0	Caucasian
0	0	Hispanic

Marker	# of Mutations
DYF399S1	15
DYF403S1a/b	11
DYS627	7
DYS612	7
DYS518	6
DYS570	5
DYS626	5
DYS547	4
DYS526a/b	3
DYS576	3
DYS449	3
DYF404S1	1
DYF387S1	1



Summary

- Rapidly Mutating Y-STRs are highly diverse markers that can discriminate common haplotypes and close relatives.
- These markers may create interpretational issues for paternity/missing persons cases.
- An international consortium is gathering frequency and mutation rate data.

Acknowledgments

NIST Team for This Work

Group Leader



**John
Butler**

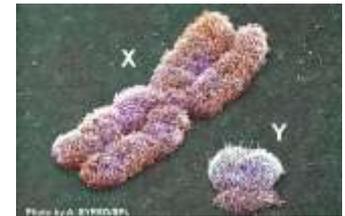


**Becky
Hill**

RM-YSTR Consortium



Manfred Kayser



Arwin Ralf

Contact Info:
mcoble@nist.gov
301-975-4330

Funding from the
**National Institute of
Justice (NIJ)** through
NIST Office of Law
Enforcement Standards

NIST Disclaimer: Certain commercial equipment, instruments and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or it imply that any of the materials, instruments or equipment identified are necessarily the best available for the purpose.

Points of view are those of the presenters and do not necessarily represent the official position of the National Institute of Standards and Technology or the U.S. Department of Justice.